
La transcription perceptuelle au service du corpus de conversations naturelles

Perceptual Transcription for Use in Natural Conversations Corpus

Geneviève Pinard-Prevost



Édition électronique

URL : <http://journals.openedition.org/praxematique/1184>

DOI : 10.4000/praxematique.1184

ISSN : 2111-5044

Éditeur

Presses universitaires de la Méditerranée

Édition imprimée

Date de publication : 1 janvier 2010

Pagination : 351-366

ISBN : 978-2-36781-012-6

ISSN : 0765-4944

Référence électronique

Geneviève Pinard-Prevost, « La transcription perceptuelle au service du corpus de conversations naturelles », *Cahiers de praxématique* [En ligne], 54-55 | 2010, document 20, mis en ligne le 01 janvier 2013, consulté le 25 septembre 2020. URL : <http://journals.openedition.org/praxematique/1184> ; DOI : <https://doi.org/10.4000/praxematique.1184>

Geneviève Pinard-Prevost
CATIFQ, université de Sherbrooke (Québec, Canada)

La transcription perceptuelle au service du corpus de conversations naturelles

Introduction

Les enregistrements de conversations naturelles entre proches, au sens strict proposé notamment par Traverso¹ (1996 : 5), recèlent de riches informations linguistiques qui se font plus rares dans d'autres contextes. Les données primaires issues de tels enregistrements sont spontanées, ni scriptées ni planifiées d'avance, et la notion d'« erreur » y perd son pouvoir discriminatoire ; cela ouvre la voie à une description mieux alignée sur la réalité orale. En négliger la transcription, particulièrement en ce qui concerne la prosodie et les gestes, c'est d'abord rendre ces matériaux de la communication difficiles d'accès à l'analyse, mais c'est aussi réduire la possibilité de mieux comprendre les enjeux verbaux dans la conversation.

Or, dans ce contexte, les voix des locuteurs sont fréquemment confondues en un même signal sonore lors des chevauchements de parole et des conversations parallèles. Toutes ces situations compromettent l'analyse informatique de la prosodie, car les logiciels disponibles à ce jour ne permettent pas l'analyse séparée des signaux confondus.

Dans notre corpus, nous avons calculé qu'environ 65 % du temps d'enregistrement était en signal vocal pur, donc le fait d'un seul locuteur. Cependant, il appert que la moitié des indices de la prosodie relevés dans les transcriptions qui en résultent marquent des syllabes produites en signal mixte. C'est donc dire qu'une plus grande proportion d'indices prosodiques se manifeste pendant les chevauchements

1. L'échange se déroule en contexte non manipulé où chaque participant a accès à la position de locuteur ; l'alternance des tours de parole n'est pas prédéterminée ; l'utilité immédiate de la conversation est le plaisir, le jeu ou la politesse ; l'informel et le « léger » y prédominent ; les relations entre les locuteurs s'y manifestent, ainsi que certains niveaux de complicité.

de parole et les conversations parallèles (50 % des indices dans 35 % de l'enregistrement), d'où l'importance d'étudier ces occurrences plutôt que de les éluder à cause des limites technologiques.

Lorsqu'il s'agit d'analyser la prosodie pour elle-même, la solution de l'enregistrement en chambre sourde suffit peut-être à solutionner le problème. Cependant, lorsque l'intérêt porte sur la prosodie en tant que phénomène interactif constitutif des conversations naturelles, l'option de l'enregistrement en laboratoire est inopportune.

Ainsi, il nous semble que la transcription perceptuelle, qui ne fait appel qu'à l'oreille du transcripteur, soit la seule avenue possible malgré tous les problèmes que cela suppose, illustrés notamment dans Pallaud (2002). Notre réflexion s'articule autour de trois qualités essentielles des transcriptions, soulignées par Mondada (2004) : la pertinence des indices transcrits ¹ en regard de la finalité du corpus, leur fidélité ² aux données primaires, et la lisibilité de la transcription.

Après avoir décrit la prosodie en interaction comme nous l'envisageons dans cet article, nous détaillerons notre méthodologie. Nos résultats nous mèneront ensuite vers quelques propositions pour permettre une transcription pertinente, puis fidèle, et lisible pour des analyses interactionnelles.

1. Objet de l'étude

Nous nous intéressons à la prosodie en tant que composante de l'interaction, dont il est utile d'inclure les manifestations dans toute transcription de l'oral destinée à l'analyse multimodale (Mertens 2008).

1.1. La prosodie

La prosodie s'analyse à travers quatre paramètres de la voix. Le tableau ci-dessous résume le lien entre leur perception et leur produc-

1. Pour les besoins de cette recherche, sont pertinents les indices de la prosodie qui, lors de leur lecture dans la transcription, aident à l'interprétation de l'interaction pour des recherches lexico-sémantiques et pragmatique de type interactionnel (et non pas phonétiques ou phonologiques).

2. Une transcription fidèle des indices prosodiques en fournit la nature et la valeur exactes, telles que produites lors de la conversation. Nous avons validé ces données acoustiquement à l'aide du logiciel Praat.

Tableau 1. — Les quatre indices de la prosodie

Paramètres perceptibles	Nom de l'indice	Corrélat acoustique
voix grave ou aiguë	hauteur	vitesse de vibration des cordes vocales
voix douce ou forte	intensité	pression sous-glottique
allongement des sons (relatif aux sons voisins)	durée	durée du son
pauses (énoncé fluide ou tronqué)	pause	formes et durées des pauses

tion, ainsi que le nom que nous utiliserons pour les désigner dans la présente étude. Ce choix repose sur un examen préalable de la terminologie utilisée pour dénommer les indices de la prosodie (Pinard-Prévost 2009).

1.2. L'approche interactionnelle de la prosodie

Portes (2004) consacre tout un chapitre de sa thèse de doctorat à la prosodie en discours. Dans cette perspective, elle signale que celle-ci jouerait les quatre rôles que voici :

- 1 Une fonction de structuration de la parole ;
- 2 Une fonction d'identification du locuteur ;
- 3 Une fonction d'expression de l'état émotionnel du locuteur et de son attitude vis-à-vis de la conversation en construction ;
- 4 Diverses fonctions interactionnelles, à savoir :
 - a La mise en valeur de l'information importante ;
 - b La gestion des tours de parole ;
 - c La contextualisation, c'est-à-dire l'adaptation de la parole à la situation de communication.

Pour représenter ce rôle en intercompréhension, la transcription des indices prosodiques et de leur valeur relative suffit.

2. Hypothèses de travail

En partant de l'idée que, lors d'une conversation, les locuteurs sont en mesure de percevoir correctement et d'interpréter de manière adéquate les signaux prosodiques produits par leurs interlocuteurs, nous posons une première hypothèse :

Le transcripteur peut, sans support informatique, percevoir adéquatement les variations prosodiques pertinentes pour l'analyse lexicale, sémantique et pragmatico-interactionnelle des conversations.

Une fois relevées ces variations prosodiques pertinentes, le transcripteur devra subséquemment déterminer la nature exacte des indices en jeu dans les saillances, afin de les transcrire précisément selon une convention de transcription donnée. Cette étape rend disponibles, sur le support tangible de la transcription écrite, les phénomènes prosodiques produits. À ce sujet, nous posons une seconde hypothèse : Le transcripteur peut distinguer entre eux, avec fidélité, les différents indices de la prosodie qui composent les saillances prosodiques.

3. Méthodologie

Nous avons conduit notre recherche en utilisant des extraits d'enregistrement issus du Corpus de français parlé au Québec (C.F.P.Q.) accompagnés de leur transcription. Le but escompté n'était pas d'analyser la qualité de la transcription dans le cadre précis de ce corpus, mais bien de cerner les limites des analyses perceptuelle et informatique d'un corpus constitué exclusivement d'enregistrements de conversations où un travail substantiel a déjà été fait pour la transcription de la prosodie. Ce corpus, élaboré depuis 2006 par la professeure Gaétane Dostie à l'université de Sherbrooke, a pour objectif de « refléter le français québécois en usage dans les années 2000 » (C.F.P.Q. 2006). Au moment d'écrire ces lignes, un peu plus de 27 heures de transcription sont disponibles en ligne, et environ 18 heures d'enregistrement supplémentaires sont soit en cours de transcription, soit en attente de l'être. Le projet total prévoit de rendre disponibles les transcriptions intégrales de 60 heures d'enregistrement. Les conventions de transcription adoptées dans le cadre de ce projet prévoit notamment les marques suivantes concernant la prosodie :

Tableau 2. — Conventions de transcription de la prosodie au C.F.P.Q.

	montée	forte montée	descente	forte descente
hauteur (hz)	/	↑	\	↓
intensité (dB)	<f<...>>	<ff<...>>	<p<...>>	<pp<...>>
accent d'intensité	MAJ			
	rapide	lent	croissant	décroissant
intensité (dB)	n/a	n/a	<cresc<...>>	<dim<...>>
débit de parole	<all<...>>	<len<...>>	<acc<...>>	<rall<...>>
allongement	: ou :: ou :::			
pauses	brèves : (.) < 1 seconde		longues : (r'') > 1 seconde	
schwas	euh			
multi-transcription	{A;B}			
discours rapporté direct	• ... °			

Nous avons sélectionné trente minutes d'extraits parmi des enregistrements mettant en scène quatre locuteurs qui discutent librement dans un lieu familier à tous. Nous avons soumis ces extraits à l'oreille de trois transcrip-teurs expérimentés du C.F.P.Q., des étudiants aux cycles supérieurs en linguistique non experts en prosodie. Ils avaient pour mandat de transcrire les proéminences prosodiques telles qu'ils les percevaient, à l'aide des conventions de transcription citées. Nous avons ensuite procédé à une analyse acoustique fine d'un extrait de quatre minutes¹, à l'aide du logiciel de traitement du signal vocal Praat, pour ensuite comparer l'analyse perceptuelle à l'analyse informatique. Précisons que, si l'analyse acoustique est informatisée, elle débute néanmoins par un découpage manuel en mots, syllabes ou phonèmes. Nous ne connaissons aucun moyen de faire l'économie de cette étape perceptuelle. Martin résume la difficulté : « Il ne faut pas perdre de vue non plus que la segmentation ne peut être qu'approchée, puisqu'il n'y a pas de frontière physique précise dans le signal qui corresponde aux sons de la parole tels que perçus par l'auditeur ni bien sûr aux phonèmes (entités formelles abstraites) puisque la parole résulte d'un geste articulatoire continu » (2008 : 97). L'exécution de cette première tâche

1. La transcription de cet extrait est disponible sur le site Internet du C.F.P.Q. (sous-corpus 4, segment 6, de la page 67, ligne 13 à la page 72, ligne 13).

demande un peu plus d'une heure trente de travail par minute de signal vocal.

Une première lecture de la prosodie dans ces transcriptions perceptuelles nous a permis d'observer, tout comme Simon, que « les trois grandes organisations prosodiques que sont l'accentuation, l'intonation et le rythme, s'ancrent sur les syllabes proéminentes » (2007 : 11). C'est donc sur les 844 syllabes de notre extrait que nous avons mesuré les valeurs minimales, maximales et moyennes de hauteur (Hz) et d'intensité (dB) de la voix, ainsi que leur durée (ms), pour ensuite repérer les syllabes saillantes. Toute la question de la pertinence se situe au niveau des seuils de proéminence.

La difficulté réside dans le fait que l'analyse prosodique de la conversation se base sur les mouvements proéminents de la prosodie, alors que celle-ci participe de la moindre manifestation vocale. Le locuteur superpose sa prosodie à ses mots, de façon tout à fait subjective (volontaire ou non), et l'interlocuteur fait illico l'interprétation de ce fil prosodique. Caelen-Haumont décrit clairement l'importance de ce *continuum* prosodique :

Ainsi au sein d'une échelle subjective, tel mot au relief mélodique plat jouant au regard du locuteur un rôle marginal dans l'expression du sens, tel autre mot au relief mélodique plus accentué exerçant un rôle moins secondaire, tel autre au relief culminant assumant un rôle essentiel, la sélection des unités établie subjectivement sur une base discriminative, remplirait la fonction élective sur les plans linguistique, pragmatique, mélodique, en supportant la conviction la plus profonde (parfois involontaire ou inconsciente) de l'individu. (2005 : 193)

Certaines recherches se sont penchées sur les seuils de perception (notamment Martin 2008 et Rossi 1999). Nous avons comparé ces seuils théoriques à ceux des syllabes qu'au moins deux transpositeurs ont relevées comme étant proéminentes dans notre corpus, pour en conclure que les seuils perceptuels obtenus étaient similaires à ceux lus. Le tableau suivant compare les seuils de proéminence théoriques à ceux de nos transpositeurs, lorsqu'il y a lieu.

4. Résultats

Nous avons lu nos résultats sur la base des trois cas de figure suivants :

Tableau 3. — Seuil de proéminence théoriques et perceptuels

	valeur absolue	écart aux unités précédentes
hauteurs (en demi-tons)	n/a	théorie : 3 demi-tons
		transcripteurs en présence d'un autre indice : 3 dt seul : 7 demi-tons
intensité	n/a	théorie : 10 dB (perception du double de volume)
		transcripteurs : en présence d'un autre indice : 5 dB seul 14 dB
allongement (données identiques en théorie et chez nos transcripteurs)	200 % de la moyenne de durée des unités d'un locuteur	200 % de la durée de l'unité précédente
schwas	théorie : 300 ms	n/a
	transcripteurs : 135 ms	
pauses	théorie : 200 ms	n/a
	transcripteurs : 115 ms	

- Les proéminences relevées à la fois par les transcripteurs et Praat ;
- Les proéminences relevées par Praat seulement ;
- Les proéminences relevées par les transcripteurs seulement.

4.1. La pertinence

Les proéminences relevées tant par les transcripteurs que par Praat ont été considérées d’emblée comme pertinentes étant donné que « [l’]on peut considérer que “le linguistiquement pertinent” est intimement lié à la signification du message, et donc à son interprétation par le transcrip-teur » (Delais-Roussarie et Yoo 2011 : 16).

Quant aux proéminences relevées par l’une ou l’autre des analyses seulement, elles ont été observées de près, en regard des fonctions de la prosodie dans l’approche interactionnelle. Nous nous sommes posé la même question, très simple, à laquelle un transcrip-teur doit répondre pour décider s’il transcrira ou non une manifestation prosodi-que perçue faiblement ou de façon ambiguë : est-ce que l’absence ou la présence de cet indice provoque un changement d’interprétation de l’énoncé dans la transcription ? Dans l’affirmative, nous avons consi-déré cet indice comme pertinent. Le tableau suivant présente, dans l’absolu, le nombre de marques prosodiques jugées pertinentes ou non, et transcrites ou non.

Tableau 4. — Pertinence et transcription des proéminences

proéminences :	pertinentes annotées	pertinentes non annotées (Praat seulement)	non pertinentes annotées (1 transcripteur)
par hauteur de la voix	25	37	10
par intensité de la voix	4	1	3
par allongement de syllabe	9	4	3
schwas	14	5	1
pauses	60	8	0
total	112	55	17

Nous concluons, à partir de ces chiffres, que le marquage des schwas et des pauses a fait l'objet d'une grande unanimité entre les transcrip-teurs et l'analyse acoustique. Si le marquage de l'intensité de la voix et de l'allongement des syllabes proéminentes est moins certain, celui de la hauteur de la voix pose problème : bien que les seuils de proémi-nence retenus soient ceux des transcrip-teurs eux-mêmes, le nombre de saillances pertinentes relevées seulement par l'analyse acoustique dépasse largement celui des saillances pertinentes transcrites. La figure suivante illustre ces mêmes résultats, indice par indice.

Notre première hypothèse voulait que les transcrip-teurs puissent perce-voir, de façon pertinente, les indices de la prosodie produits en conver-sation. Nous pouvons conclure, à partir du court extrait analysé, que peu de saillances relevées par un transcrip-teur sont non pertinentes (9 %), mais que l'analyse gagnerait à être affinée. En effet, les 30 % des proéminences pertinentes non transcrites (principalement des indices de hauteur de la voix), relevées uniquement par Praat, correspondent aux niveaux de sensibilité établis à partir des performances des trans-crip-teurs eux-mêmes. Ainsi, à la question « les transcriptions percep-tuelles sont-elles pertinentes ? » nous pouvons répondre par l'affirma-tive. Cependant, à la question « les transcrip-teurs ont-ils pu percevoir

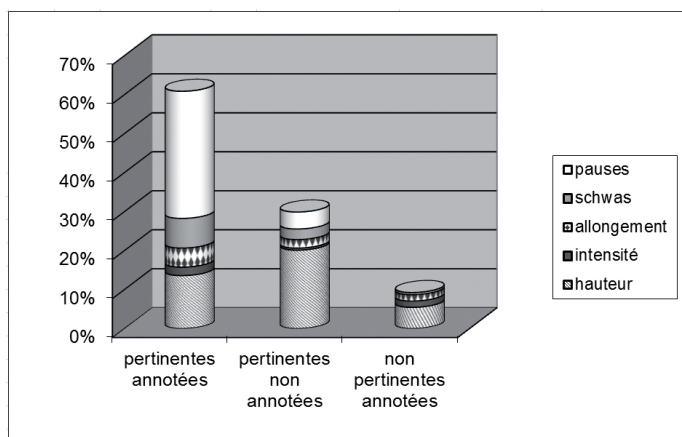


Figure 1. — Transcription (ou non) des manifestations pertinentes de la prosodie.

toutes les saillances pertinentes? » la réponse est plus nuancée : la sensibilité de l'analyse perceptuelle gagnerait à être améliorée.

4.2. La fidélité

Afin de déterminer le niveau de fidélité de la transcription perceptuelle aux données primaires, nous avons comparé l'indice choisi pour marquer la proéminence par l'une et l'autre des analyses.

Nous avons constaté que la plupart des problèmes de fidélité survenaient sur les syllabes où l'analyse acoustique a relevé plusieurs indices saillants simultanés : des combinaisons de mouvement de hauteur de la voix doublé d'une forte intensité ou d'un allongement. Les transcriviteurs, en accord avec la convention de transcription du C.F.P.Q., ont signifié ces proéminences en transcrivant les syllabes pertinentes en lettres majuscules (MAJ).

Nous avons également observé que la transcription des mouvements de hauteur (/ et ↑; \et ↓) était peu cohérente. Nos résultats laissent entrevoir qu'il devient malaisé de fixer un seuil pour distinguer les montées ou descentes de hauteur moyennes « / \ » de celles qui sont fortes « ↑ ↓ ». Cependant, il nous semble essentiel de transcrire dif-

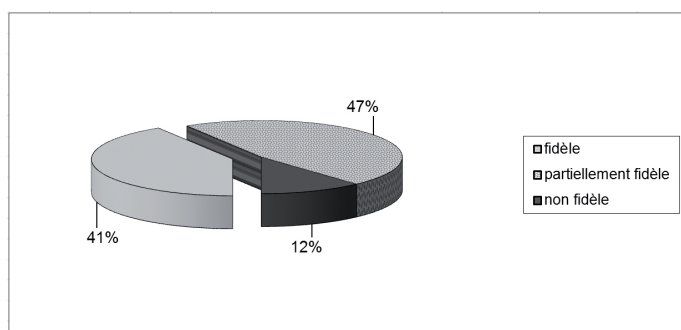


Figure 2. — Proportion des différents niveaux de fidélité.

féremment les montées de hauteur à interprétation interrogative et celles qui ne le sont pas, interprétation que seule l'analyse perceptuelle peut faire. Mentionnons d'ores et déjà qu'il en est de même pour le repérage du discours rapporté direct : en effet, un logiciel ne peut pas décider si telle ou telle variation de hauteur de la voix témoigne de la personnification d'une voix narrative externe à celle du locuteur, alors qu'un transcrip-teur peut, lui, indiquer cette interprétation dans sa transcription. La convention de transcription du C.F.P.Q. prévoit l'utilisation d'un point noir d'ouverture et d'un point blanc de fermeture pour en signifier la présence.

La figure suivante illustre la proportion de transcriptions fidèles (correspondance exacte entre la transcription perceptuelle et l'analyse acoustique), partiellement fidèles (mêmes indices, mais valeurs ou mouvements différents) et non fidèles (indices différents ou MAJ).

Notre conclusion concernant la fidélité de la transcription perceptuelle ne peut donc être que la suivante : certes, un transcrip-teur peut déterminer avec justesse la nature d'un indice de la prosodie qui agit comme « surligneur » sur une syllabe proéminente, mais il en jauge plus difficilement la valeur exacte.

5. Pour une transcription perceptuelle pertinente, fidèle et lisible

Notre problématique de la pertinence d'une transcription touche des manifestations prosodiques en lien avec la structuration de l'énoncé, la mise en valeur de l'information importante et la gestion des tours de parole. La question de la fidélité concerne plutôt le raffinement de cette transcription, ce qui inclut des indices qui, selon Portes (2004), concernent le marquage de la fonction d'identification du locuteur, de l'expression de son état émotionnel, de son attitude vis-à-vis de la conversation, et de la contextualisation.

5.1. Une transcription prosodiquement pertinente

Notre premier objectif est l'augmentation du degré de pertinence de la transcription, en suggérant au transcripteur de se concentrer sur les effets prosodiques (structuration, interrogations, mises en évidence), plutôt que sur les indices. Ainsi, le relevé prosodique initial sera plus précis quant à son interprétation, quoique moins exact quant à la nature des indices et leurs valeurs, étant donnée l'utilisation des MAJ. Voici, de manière plus concrète, nos propositions :

- Transcrire d'abord les pauses, les schwas et le discours rapporté direct en même temps que le matériel verbal ;
- Transcrire également les mouvements de hauteur à valeur illocutoire de question par le symbole « ↑ » ;
- Transcrire en majuscules (MAJ) les mots qui semblent avoir été mis en évidence par les locuteurs, sans chercher à faire une analyse plus exacte de la nature et de la valeur des indices en jeu ;
- Indiquer les syllabes allongées par un deux-points.

Cette proposition fait écho aux conclusions de Delais-Roussarie et Yoo : « les symboles [pour la transcription de la prosodie] doivent rendre compte de l'accentuation, de l'intonation et des groupements prosodiques, et cela en tentant de rester le plus neutre possible sur le plan théorique [...] les unités auxquelles les symboles choisis seront assignés peuvent difficilement être les mêmes pour traiter de l'accentuation et de l'intonation » (2011 : 20).

Nous avons sollicité nos trois transpositeurs pour une validation des étapes précitées. En voici les résultats :

- 100 % des très forts mouvements de hauteur à valeur illocutoire de question ont été transcrits ;
- 77 % des syllabes proéminentes par un mouvement de hauteur ou d'intensité ont été marquées ;
- 72 % des syllabes allongées ont été marquées.

Ces résultats sont meilleurs que ceux obtenus lors de la première analyse où la perception des saillances et l'identification des indices étaient confondues, alors que 30 % des proéminences pertinentes n'avaient pas été relevées par les transpositeurs. Nous pensons que des résultats encore plus satisfaisants seraient obtenus en contexte réel de transcription, à savoir un travail d'équipe permettant la concertation, tel qu'exécuté au C.F.P.Q.

5.2. Une transcription fidèle aux données primaires

Pour une deuxième phase de transcription, nous suggérons les étapes suivantes, directement inspirées par la convention de transcription du C.F.P.Q. :

1. Reprendre les syllabes proéminentes MAJ et transcrire fidèlement leurs mouvements de hauteur (*/*) ou d'intensité forte <f>, selon le cas ;
2. Transcrire les extraits criés en utilisant *fortissimo* <ff> ;
3. Transcrire les voix basses par *piano* <p> et les voix très basses ou les chuchotements par *pianissimo* <pp> ;
4. Transcrire les extraits qui semblent lents par *lento* <len> et ceux qui semblent rapides par *allegro* <all>.

Les transpositeurs ont l'avantage de percevoir les phénomènes prosodiques dans les segments en chevauchement de parole ou en conversation parallèle, ce qu'aucun outil informatique n'arrive encore à faire. Nous avons également relevé des cas probants d'accélération du débit de parole qui n'ont pu être confirmés par l'analyse acoustique, mais dont l'effet semble indéniable. Il nous semble important que ce phénomène soit rendu dans la transcription. Même si les tracés de Praat ne permettent pas de visualiser cette accélération, il n'en demeure pas

moins que cet effet est nettement audible, et qu'elle pourrait avoir un impact sur l'interaction. Évidemment, là n'est pas l'objet de notre étude, mais cet autre phénomène mériterait investigation, d'où l'importance de sa présence dans la transcription. En effet, comment un chercheur pourrait-il avoir la puce à l'oreille et s'intéresser à ce genre de cas s'ils sont oblitérés dans la transcription, faute de « preuves » acoustiques ? La preuve ne demande peut-être qu'à être construite autrement.

5.3. Une transcription « interactionnellement » lisible

L'obligatoire lecture linéaire de la transcription rend assez peu justice au dynamisme de la conversation. Il demeure difficile, pour le lecteur, de se représenter l'interaction, notamment lorsque plusieurs très courts chevauchements interviennent dans le monologue d'un locuteur.

Nous envisageons trois avenues pour rendre « interactionnellement » lisible la transcription d'une conversation :

1. Transcrire toutes les pauses produites, même celles qui donnent lieu à un changement de locuteur ;
2. Délimiter les chevauchements de parole symétriquement et les relier par des traits ;
3. Transcrire côte à côte les conversations parallèles.

La transcription des pauses qui donnent lieu à un changement de locuteur (et donc à un changement de ligne dans la transcription) nous semble importante, car elles pourraient véhiculer certaines informations pertinentes à propos du legs du tour de parole. Les énoncés qui s'achèvent par une pause de ce genre diffèrent grandement des cas de chevauchement où un interlocuteur s'insère dans la conversation sans qu'aucune pause ne lui ait préalablement pratiqué une ouverture.

La transcription d'un chevauchement de parole illustre difficilement l'équivalence de durée des deux segments simultanés. Pour y remédier, nous proposons de disposer symétriquement les paires de crochets qui délimitent les chevauchements et d'y centrer les énoncés graphiquement plus courts afin de démontrer les durées équivalentes.

Voici, ci-après, une illustration du résultat obtenu à la suite de l'application de ces quelques conventions graphiques à un extrait de transcription du C.F.P.Q.

M = ben il y en avait/ qui étaient bons/ aussi (.) {parce que } j'en avais fait euh chacun
aux enfants les nappes et[pis] (.) pis ils les ont gardées ben longtemps (.) pis ils
mettaient tout le temps ceux-là (.) {(RIRE)}

H = [oui]

H = [hum] (en hochant la tête affirmativement)

A = [mais] le matériel/ était bon/ le (.) c'est[
{parce euh le le} ça a été mal euh (.)[étiré] le (.) c'est[

M = [oui]

H = [oui]

H = [ajusté là ouin ouin] {ouin;O}

M = fait que ceux-là ben il disait •jette-les° ben j'ai dit •<<mon>>doux seigneur quand
même ce sera [croche un peu (.)] on mange pareil°

H = [(RIRE)] BEN: oui (.) oui (.)

Figure 3. — Transcription des chevauchements.

L'extrait de transcription est disponible sur le site du C.F.P.Q. (sous-corpus 4, segment 6, page 69, lignes 9 à 17).

Enfin, nous proposons de présenter les conversations parallèles en encadrant chacune d'elle et en les plaçant côte à côte. Nous croyons que ce parallélisme physique représente mieux la simultanéité des deux conversations, comme en atteste l'illustration suivante :

A = il était au foyer / ah \

M = parce que: elle pouvait plus le garder parce qu'elle euh (.) franchement euh	R = à Sainte-Marguerite (2.4'') t'sais en arrière de: l'église Sainte-Marguerite là
H = ben elle est pas mal plus jeune euh: que lui hein† (.) ouais (en hochant la tête affirmativement)	A = oui oui (.) oui c'est beau c'est beau ce foyer-là
M = elle a soixante-dix-neuf	R = ah je sais pas je suis jamais allé moi
H = soixante-dix-neuf†	A = très beau
M = ouais (en hochant la tête affirmativement)	R = qu'est-ce qu'ils vont faire avec l'église à Sainte-Marguerite

A = probablement l'église une partie là les il faut qu'ils défassent les clochers ça c'est
certain (.) sont après tomber (1.5'') puis euh: pis c'est pas vieux hein ça cette affaire-là
là [c'est parce que ça a été mal bâti

Figure 4. — Transcription de conversations parallèles.

L'extrait de transcription est disponible sur le site du C.F.P.Q. (sous-corpus 4, segment 6, page 50, lignes 3 à 15).

Conclusion

Après vérification de notre première hypothèse relative à la transcription perceptuelle pertinente, il appert qu'une proportion non négligeable d'un peu plus de 20 % de saillances sont passées sous silence, malgré un recensement simplifié des proéminences. Nous savons cependant que la moitié des indices de la prosodie participent d'un chevauchement de parole ou de conversations parallèles, donc demeurent inanalysables informatiquement... Entre 50 % d'indices sacrifiés par l'analyse acoustique ou 20 % par l'analyse perceptuelle, le choix nous semble évident.

Notre deuxième hypothèse stipulait que le transcripteur pouvait également, à l'oreille, distinguer entre eux les différents indices de la prosodie et les transcrire fidèlement par rapport aux données premières. Le premier pari a été tenu, mais non le second. Une transcription perceptuelle peut rendre avec précision la nature des indices prosodiques en présence (hauteur et intensité de la voix, durée et pauses), mais non leur valeur exacte.

À partir de ces conclusions, nous avons proposé une transcription qui rende lisible le dynamisme des interactions.

Évidemment, notre recherche, dans son ensemble, prend appui sur des conversations spontanées où les choix initiaux qui ont présidé à la constitution du corpus favorisent un contexte familier, au détriment de la pureté du signal vocal : d'autres types de corpus mèneraient vraisemblablement à d'autres conclusions. Cela dit, il nous semble que la conversation la plus spontanée possible constitue un matériau incontournable pour mener des recherches lexico-sémantiques et pragmatico-interactionnelles, même si cela nous oblige à faire des choix (voire des sacrifices) méthodologiques.

Plusieurs explorations très intéressantes restent encore à faire sur ce riche patrimoine. Jusqu'à quel point doit-on négliger les nuances inhérentes à la conversation, au nom d'une rigueur technologique tout simplement peu adaptée à nos besoins d'analyse (et à nos aspirations)? Nous laissons la question en suspens.

Références bibliographiques

CAELEN-HAUMONT G.,

2005, *Prosodie et sens : une approche expérimentale*, Paris, L'Harmattan.

- C.F.P.Q., DOSTIE G. *et al.*,
2006, Catifq, Université de Sherbrooke, [En ligne] : <http://pages.usherbrooke.ca/cfpq/>.
- DELAIS-ROUSSARIE É. & YOO H.-Y.,
2011, « Transcrire la prosodie : un préalable à l'échange et à l'analyse de données », *Journal of French Language Studies* 21, 1, Cambridge University Press.
- MARTIN P.,
2008, *Phonétique acoustique : Introduction à l'analyse acoustique de la parole*, Paris, Armand Colin.
- MERTENS P.,
2008, « Syntaxe, prosodie et structure informationnelle : une approche prédictive pour l'analyse de l'intonation dans le discours », *Travaux de linguistique* 1, 56, 97-124.
- MONDADA L.,
2004, « Approche des corpus en linguistique interactionnelle », université Stendhal Grenoble 3, [En ligne] : <http://w3.u-grenoble3.fr/lebarbe/elc/supports/mondada.pdf>.
- PALLAUD B.,
2002, « Erreurs d'écoute dans la transcription de données orales », *Revue Parole* 22-23-24, [En ligne] : <http://halshs.archives-ouvertes.fr/docs/00/26/51/94/PDF/1405.pdf>.
- PINARD-PRÉVOST G.,
2009, « Un consensus terminologique en prosodie? », in FLORENTIN V. & SOW F. (dir.), Actes des XXIII^e Journées de la linguistique (J.D.L.), 5-6 mars 2009, université Laval, Québec, CIRAL.
- PORTES C.,
2004, *Prosodie et économie du discours : spécificité phonétique, écologie discursive et portée pragmatique du patron d'implication*, thèse de doctorat, sous la direction de Albert Di Cristo, université de Provence.
- ROSSI M.,
1999, *L'intonation, le système du français : description et modélisation*, Paris, Ophrys, coll. « L'essentiel français ».
- SIMON A.-C.,
2007, « Guide méthodologique : transcription outillée — prosodies », École thématique du C.N.R.S., Lyon, 2007, [En ligne] : http://icar.univ-lyon2.fr/ecole_thematique/contaci/.
- TRAVERSO V.,
1996, *La conversation familière. Analyse pragmatique des interactions*, Lyon, Presses universitaires de Lyon.